

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

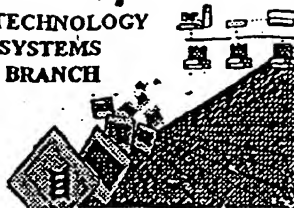
- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**



BIOTECHNOLOGY
SYSTEMS
BRANCH



RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/605,708
Source: IFWO
Date Processed by STIC: 3/25/04

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221

Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/efb/efs/downloads/documents.htm>), EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER: 10/605,708

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 ☐ Wrapped Nucleics
Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 ☒ Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 ☒ Misaligned Amino
Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 ☐ Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 ☐ Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 ☐ PatentIn 2.0
"bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s). Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 ☐ Skipped Sequences
(OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 ☐ Skipped Sequences
(NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000
- 9 ☐ Use of n's or Xaa's
(NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 ☐ Invalid <213>
Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 ☐ Use of <220> Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses.
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 01/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 ☐ PatentIn 2.0
"bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 ☐ Misuse of n/Xaa "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid



IFWO

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/605,708

DATE: 03/25/2004

TIME: 09:07:31

Input Set : D:\GLOF007USC1.txt

Output Set: N:\CRF4\03252004\J605708.raw

1 <110> APPLICANT: GONG, Zhiyuan
2 LAM, Toong Jin
3 JU, Bensheng
4 XU, Yanfei
5 HE, Jiangyan
6 YAN, Tie
8 <120> TITLE OF INVENTION: CHIMERIC GENE CONSTRUCTS FOR GENERATION OF
9 FLUORESCENT TRANSGENIC ORNAMENTAL FISH
11 <130> FILE REFERENCE: 1781-0163P
13 <140> CURRENT APPLICATION NUMBER: US/10/605,708
14 <141> CURRENT FILING DATE: 2003-10-21
16 <160> NUMBER OF SEQ ID NOS: 24
17 <170> SOFTWARE: PatentIn Ver. 2.0

ERRORED SEQUENCES

21 <210> SEQ ID NO: 1
22 <211> LENGTH: 2480
23 <212> TYPE: DNA
24 <213> ORGANISM: Danio rerio
26 <220> FEATURE:
27 <221> NAME/KEY: CDS
28 <222> LOCATION: (90)..(1586)
30 <220> FEATURE:
31 <221> NAME/KEY: primer bind
32 <222> LOCATION: (66)..(85)
33 <223> OTHER INFORMATION: CK2
35 <220> FEATURE:
36 <221> NAME/KEY: primer bind
37 <222> LOCATION: (97)..(120)
38 <223> OTHER INFORMATION: CK1
40 <220> FEATURE:
41 <221> NAME/KEY: polyA signal
42 <222> LOCATION: (24467)..(2451)
44 <400> SEQUENCE: 1
45 ctctcctttg tgagcaacct cctccactca ctctctctc agagagcact ctctgacctc 60
46 ctctcagca actcaaagac acaggcatc atg tca acc agg tct atc tct tac 113
47 Met Ser Thr Arg Ser Ile Ser Tyr
48 1 5
50 tcc agc ggt ggc tcc atc agg agg ggc tac acc agc cag tca gcc tat 161
51 Ser Ser Gly Gly Ser Ile Arg Gly Tyr Thr Ser Gln Ser Ala Tyr
52 10 15 20 20

Does Not Comply
Corrected Diskette Needed
(ps. 1-5)

please see item #
3 on error summary
sheet.

RAW SEQUENCE LISTING

DATE: 03/25/2004

PATENT APPLICATION: US/10/605,708

TIME: 09:07:31

Input Set : D:\GLOF007USC1.txt

Output Set: N:\CRF4\03252004\J605708.raw

```

54 gca gta cct gcc ggc tct acc agg atg agc tca gtg acc agt gtc agg 209
55 Ala Val Pro Ala Gly Ser Thr Arg Met Ser Ser Val Thr Ser Val Arg
--> 56 25 30 35 40
58 aga tct ggt gtg ggt gcc agc cca ggc ttc ggt gcc ggt ggc agc tac 257
59 Arg Ser Gly Val Gly Ala Ser Pro Gly Phe Gly Ala Gly Gly Ser Tyr
--> 60 45 50 55
62 agc ttt agc agc agc agc atg ggt gga ggc tat gga agt ggt ctt ggt 305
63 Ser Phe Ser Ser Ser Ser Met Gly Gly Gly Tyr Gly Ser Gly Leu Gly
--> 64 gga ggt ctc ggt ggt ggc atg ggc ctt cgt tgc ggg ctt ctt atc aca 353
65 Gly Gly Leu Gly Gly Gly Met Gly Phe Arg Cys Gly Leu Pro Ile Thr
--> 66 75 80 85
69 gct gta act gtc aac cag aac ctg ttg gcc ccc tta aac ctg gaa atc 401
70 Ala Val Thr Val Asn Gln Asn Leu Leu Ala Pro Leu Asn Leu Glu Ile
--> 71 90 95 100
73 gac ccc aca att caa gtc cgc act tca gag aaa gag cag att aag 449
74 Asp Pro Thr Ile Gln Ala Val Arg Thr Ser Glu Lys Glu Gln Ile Lys
--> 75 105 110 115 120
77 acc ttc aac aac cgc ttc gct ttc ctc atc gac aaa gtg cgc ttc ctg 497
78 Thr Phe Asn Asn Arg Phe Ala Phe Leu Ile Asp Lys Val Arg Phe Leu
--> 79 125 130 135
81 gaa cag cag aac aag atg ctt gag acc aaa tgg agt ctt ctc caa gaa 545
82 Glu Gln Gln Asn Lys Met Leu Glu Thr Lys Trp Ser Leu Leu Gln Glu
--> 83 140 145 150
85 cag aca acc aca cgt tcc aac atc gat gcc atg ttt gag gca tac atc 593
86 Gln Thr Thr Thr Arg Ser Asn Ile Asp Ala Met Phe Glu Ala Tyr Ile
--> 87 155 160 165
89 tct aac ctg cgc aga cag ctc gat gga ctg gga aat gag aag atg aag 641
90 Ser Asn Leu Arg Arg Gln Leu Asp Gly Leu Gly Asn Glu Lys Met Lys
--> 91 170 175 180
93 ctg gag gga gag ctg aag aac atg cag ggc ctg gtt gag gac ttc aag 689
94 Leu Glu Gly Glu Leu Lys Asn Met Gln Gly Leu Val Glu Asp Phe Lys
--> 95 185 190 195 200
97 aac aag tac gag gat gag atc aac aag cgt gct tcc gta gag aat gag 737
98 Asn Lys Tyr Glu Asp Glu Ile Asn Lys Arg Ala Ser Val Glu Asn Glu
--> 99 205 210 215
101 ttt gtc ctg ctc aag aag gat gtt gat gca gcc tac atg aac aag gtt 785
102 Phe Val Leu Leu Lys Lys Asp Val Asp Ala Ala Tyr Met Asn Lys Val
--> 103 220 225 230
105 gag ctt gaa gcc aag gtt gat gct ctt cag gat gag atc aac ttc ctc 833
106 Glu Leu Glu Ala Lys Val Asp Ala Leu Gln Asp Glu Ile Asn Phe Leu
--> 107 235 240 245
109 agg gca gtc tac gag gct gaa ctc cgg gag ctc cag tct cag atc aag 881
110 Arg Ala Val Tyr Glu Ala Glu Leu Arg Glu Leu Ser Gln Ile Lys
--> 111 250 255 260
113 gac aca tct gtt gtt gta gaa atg gac aac agc aga aac ctg gat atg 929
114 Asp Thr Ser Val Val Val Glu Met Asp Asn Ser Arg Asn Leu Asp Met
--> 115 265 270 275 280
117 gac tcc atc gtg gct gaa gtt cgc gct cag tat gaa gac atc gcc aac 977
118 Asp Ser Ile Val Ala Glu Val Arg Ala Gln Tyr Glu Asp Ile Ala Asn

```

Invalid
Nucleotide
code

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/605,708

DATE: 03/25/2004

TIME: 09:07:31

Input Set : D:\GLOF007USC1.txt

Output Set: N:\CRF4\03252004\J605708.raw

Same errors
N

```

-> 119          285          290          295
121 cgc agc cgt gcc gag gca gag agc tgg tac aaa cag aag ttt gag gag 1025
122 Arg Ser Arg Ala Glu Ala Glu Ser Trp Tyr Lys Gln Lys Phe Glu Glu.
-> 123          300 300 SAME error 305 305          310 310
125 atg cag agc acc ggt ggt cag tat ggt gat gac ctc cgc tca aca aag 1073
126 Met Gln Ser Thr Ala Glu Gln Tyr Gly Asp Asp Leu Arg Ser Thr Lys
-> 127          315 315 SAME error 320 320          325 325
129 gct gag att ggt gaa ctc aac cgc atg atc gcc cgc ctg cag aac gag 1121
130 Ala Glu Ile Ala Glu Leu Asn Arg Met Ile Ala Arg Leu Gln Asn Glu
-> 131          320 330          335 335          340 340
133 atc gat gtc aag gca cag cgt gcc aac ttg gag gct cag att gct 1169
134 Ile Asp Ala Val Lys Ala Gln Arg Ala Asn Leu Glu Ala Gln Ile Ala
-> 135          345 345 INVALID Amino Acid Code 350 350          355 355
137 gag gct gaa gag cgt gga gag ctg gca gtg aag gat gcc aag ctc cgc 1217
138 Glu Ala Glu Glu Arg Gly Glu Leu Ala Val Lys Asp Ala Lys Leu Arg
-> 139          360          365 365          370 370          375 375
141 atc agg gag ctg gag gaa get ctt cag agg gcc aag cag gac atg gcc 1265
142 Ile Arg Glu Leu Glu Glu Ala Leu Gln Arg Ala Lys Gln Asp Met Ala
-> 143          380 380          385 385          390 390
145 cgc cag gtc cgc gag tac cag gag ctc atg aac gtc aaa ttg gct ctg 1313
146 Arg Gln Val Arg Glu Tyr Gln Glu Leu Met Asn Val Lys Leu Ala Leu
-> 147          395 395          400 400          405 405
149 gac att gag atc gcc acc tac agg aaa ctg ttg gaa gga gag gag agc 1361
150 Asp Ile Glu Ile Ala Thr Tyr Arg Lys Leu Leu Glu Gly Glu Glu Ser
-> 151          410 410          415 415          420 420
153 aga ctg tcc agc ggt gga gct cag gct acc att cat gtt cag cag acc 1409
154 Arg Leu Ser Ser Gly Gly Ala Gln Ala Thr Ile His Val Gln Gln Thr
-> 155          425 425          430 430          435 435          440 440
157 tcc gga ggt gtt tca tct ggt tat ggt ggt agc ggc tct ggt ttc ggc 1457
158 Ser Gly Gly Val Ser Ser Gly Tyr Gly Gly Ser Gly Ser Gly Phe Gly
-> 159          445 445          450 450          455 455
161 tac agc agt ggc ttc agc agt ggt ggg tca gga tac ggt agt gga tca 1505
162 Tyr Ser Ser Gly Phe Ser Ser Gly Gly Ser Gly Tyr Gly Ser Gly Ser
-> 163          460 460          465 465          470 470
165 gga ttc ggt tct gga tca ggg tat ggt gga ggc tcc atc agc aaa acc 1553
166 Gly Phe Gly Ser Gly Ser Gly Tyr Gly Gly Ser Ile Ser Lys Thr
-> 167          475 475          480 480          485 485
169 agt gtc acc acc gtc agc agt aaa cgc tat taa ggagaagccc gcccaaacc 1606
170 Ser Val Thr Thr Val Ser Ser Lys Arg Tyr
-> 171          490 490          495 495
173 ccagccgaca cagtttccaa ctttccttac ctgcaactag atcccttctg aaccttctta 1666
175 cgactcaaac catctatggt gctatatatt agccagacag ctgtcccctg ttaatgagga 1726
177 gatgtggacg atgattttta aagtacaaaa taagtttttag attgttctgt gtgttgatgg 1786
179 tagttaccgc tatcatgcat ctccctgtctg gtggtgtcac tgccatttta aatcatcaac 1846
181 ccaactacac taaacagata ccaggaagaa tctgtctcca agccactgaa tagtcttatt 1906
183 tctgcactga taatgacagg gaaagtgaga cacatagaaa ccaactgtaac ctacgtagta 1966
185 ctatggtttc actggatcag ggggtgtgcta tacaagttcc tgaatgtctt gtttgaatgt 2026
187 tttgtgctgt tacaagctcc ctgctgtagt tttgctgact aatctgactt ttgtcatttt 2086
189 gctatggctg tcagagttgg tttacctatt ttctataaaa tgtatatggc agtcagccaa 2146

```

INVALID Amino Acid Code
INVALID Amino Acid Code

see item # 3 on error summary sheet.

RAW SEQUENCE LISTING

DATE: 03/25/2004

PATENT APPLICATION: US/10/605,708

TIME: 09:07:31

Input Set : D:\GLOF007USC1.txt

Output Set: N:\CRF4\03252004\J605708.raw

```

191 taactgatga caattgcttg tgggctacta atgtccagtt acctcacatt caagggagat 2206
193 ctgttacagc aaaaaacagg cacaatggga tttatgtgga ccatccctcc ttaaccttgt 2266
195 gtactttccg tggtggaagt ggtgactgta ctgccttaca cattcccctg tattcaactg 2326
197 gcttccagag catattttac atccccggtt ataaatggaa aatgcaagaa aactgaaaca 2386
199 atgttcaacc agatttattt ggtattgatt gacgagacac caacttgaaa tttgaatata 2446
201 ataaatctga gaccacaaaa aaaaaaaaaa aaaa 2480
204 <210> SEQ ID NO: 2
205 <211> LENGTH: 498 - Found 484
206 <212> TYPE: PRT
207 <213> ORGANISM: Danio rerio
209 <400> SEQUENCE: 2
210 Met Ser Thr Arg Ser Ile Ser Tyr Ser Ser Gly Gly Ser Ile Arg Arg
211 1 5 10 15
213 Gly Tyr Thr Ser Gln Ser Ala Tyr Ala Val Pro Ala Gly Ser Thr Arg
214 20 25 30
216 Met Ser Ser Val Thr Ser Val Arg Arg Ser Gly Val Gly Ala Ser Pro
217 35 40 45
219 Gly Phe Gly Ala Gly Gly Ser Tyr Ser Phe Ser Ser Ser Ser Met Gly
220 50 55 60
222 Gly Gly Tyr Gly Ser Gly Leu Gly Gly Gly Leu Gly Gly Gly Met Gly
223 65 70 75 80
225 Phe Arg Cys Gly Leu Pro Ile Thr Ala Val Thr Val Asn Asn Leu Leu
226 85 90 95
228 Ala Pro Leu Asn Leu Glu Ile Asp Pro Thr Ile Ala Val Arg Thr Ser
229 100 105 110
231 Glu Lys Glu Gln Ile Lys Thr Phe Asn Asn Arg Phe Ala Phe Leu Ile
232 115 120 125
234 Asp Lys Val Arg Phe Leu Glu Asn Lys Met Leu Glu Thr Lys Trp Ser
235 130 135 140
237 Leu Leu Glu Thr Thr Thr Arg Ser Asn Ile Asp Ala Met Phe Glu Ala
238 145 150 155 160
240 Tyr Ile Ser Asn Leu Arg Arg Leu Asp Gly Leu Gly Asn Glu Lys Met
241 165 170 175
243 Lys Leu Glu Gly Glu Leu Lys Asn Met Gln Gly Leu Val Glu Asp Phe
244 180 185 190
246 Lys Asn Lys Tyr Glu Asp Glu Ile Asn Lys Arg Ala Ser Val Glu Asn
247 195 200 205
249 Glu Phe Val Leu Leu Lys Lys Asp Val Asp Ala Ala Tyr Met Asn Lys
250 210 215 220
252 Val Glu Leu Glu Ala Lys Val Asp Ala Leu Asp Glu Ile Asn Phe Leu
253 225 230 235 240
255 Arg Ala Val Tyr Glu Ala Glu Leu Arg Glu Leu Gln Ser Ile Lys Asp
256 245 250 255
258 Thr Ser Val Val Val Glu Met Asp Asn Ser Arg Asn Leu Asp Met Asp
259 260 265 270
261 Ser Ile Val Ala Glu Val Arg Ala Gln Tyr Glu Asp Ile Ala Asn Arg
262 275 280 285
264 Ser Arg Ala Glu Ala Glu Ser Trp Tyr Lys Lys Phe Glu Glu Met Gln
265 290 295 300

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/605,708

DATE: 03/25/2004

TIME: 09:07:31

Input Set : D:\GLOF007USC1.txt

Output Set: N:\CRF4\03252004\J605708.raw

267 Ser Thr Ala Gly Gln Tyr Gly Asp Asp Leu Arg Ser Thr Lys Ala Glu
 268 305 310 315 320
 270 Ile Ala Glu Leu Asn Arg Met Ile Ala Arg Leu Asn Glu Ile Asp Ala
 271 325 330 335
 273 Val Lys Ala Arg Ala Asn Leu Glu Ala Gln Ile Ala Glu Ala Glu Glu
 274 340 345 350
 276 Arg Gly Glu Leu Ala Val Lys Asp Ala Lys Leu Arg Ile Arg Glu Leu
 277 355 360 365
 279 Glu Glu Ala Leu Gln Arg Ala Lys Asp Met Ala Arg Gln Val Arg Glu
 280 370 375 380
 282 Tyr Gln Glu Leu Met Asn Val Lys Leu Ala Leu Asp Ile Glu Ile Ala
 283 385 390 395 400
 285 Thr Tyr Arg Lys Leu Glu Gly Glu Glu Ser Arg Leu Ser Ser Gly
 286 405 410 415
 288 Gly Ala Gln Ala Thr Ile His Val Gln Thr Ser Gly Gly Val Ser Ser
 289 420 425 430
 291 Gly Tyr Gly Gly Ser Gly Ser Gly Phe Gly Tyr Ser Ser Gly Phe Ser
 292 435 440 445
 294 Ser Gly Gly Ser Gly Tyr Gly Ser Gly Ser Gly Phe Gly Ser Gly Ser
 295 450 455 460
 297 Gly Tyr Gly Gly Gly Ser Ile Ser Lys Thr Ser Val Thr Thr Val Ser
 298 465 470 475 480
 300 Ser Lys Arg Tyr - see pg. 4

611 <210> SEQ ID NO: 6

612 <211> LENGTH: 319

613 <212> TYPE: PRT

614 <213> ORGANISM: Danio rerio

616 <400> SEQUENCE: ②-⑥

617 Met Pro Arg Glu Asp Arg Ala Thr Trp Lys Ser Asn Tyr Phe Leu Lys
 618 1 5 10 15
 619 Ile Ile Gln Leu Leu Asp Asp Phe Pro Lys Cys Phe Ile Val Gly Ala
 620 20 25 30
 621 Asp Asn Val Gly Ser Lys Gln Met Gln Thr Ile Arg Leu Ser Leu Arg
 622 35 40 45
 623 Gly Lys Ala Val Val Leu Met Gly Lys Asn Thr Met Met Arg Lys Ala
 624 50 55 60
 625 Ile Arg Gly His Leu Glu Asn Asn Pro Ala Leu Glu Arg Leu Leu Pro
 626 65 70 75 80
 627 His Ile Arg Gly Asn Val Gly Phe Val Phe Thr Lys Glu Asp Leu Thr
 628 85 90 95
 629 Glu Val Arg Asp Leu Leu Leu Ala Asn Lys Val Pro Ala Ala Ala Arg
 630 100 105 110
 631 Ala Gly Ala Ile Ala Pro Cys Glu Val Thr Val Pro Ala Gln Asn Thr
 632 115 120 125
 633 Gly Leu Gly Pro Glu Lys Thr Ser Phe Phe Gln Ala Leu Gly Ile Thr
 634 130 135 140
 635 Thr Lys Ile Ser Arg Gly Thr Ile Glu Ile Leu Ser Asp Val Gln Leu
 636 145 150 155 160
 637 Ile Lys Pro Gly Asp Lys Val Gly Ala Ser Glu Ala Thr Leu Leu Asn

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/605,708

DATE: 03/25/2004

TIME: 09:07:31

Input Set : D:\GLOF007USC1.txt

Output Set: N:\CRF4\03252004\J605708.raw

```

638          165          170          175
639 Met Leu Asn Met Leu Asn Ile Ser Pro Phe Ser Tyr Gly Leu Ile Ile
640          180          185          190
641 Gln Gln Val Tyr Asp Asn Gly Ser Val Tyr Ser Pro Glu Val Leu Asp
642          195          200          205
643 Ile Thr Glu Asp Ala Leu His Lys Arg Phe Leu Lys Gly Val Arg Asn
644          210          215          220
645 Ile Ala Ser Val Cys Leu Gln Ile Gly Tyr Pro Thr Leu Ala Ser Ile
646 225          230          235          240
647 Pro His Thr Ile Ile Asn Gly Tyr Lys Arg Val Leu Ala Val Thr Val
648          245          250          255
649 Glu Thr Asp Tyr Thr Phe Pro Leu Ala Glu Lys Val Lys Ala Tyr Leu
650          260          265          270
651 Ala Asp Pro Thr Ala Phe Ala Val Ala Ala Pro Val Ala Ala Thr
652          275          280          285
653 Glu Gln Lys Ser Ala Ala Pro Ala Ala Lys Glu Glu Ala Pro Lys Glu
654          290          295          300
655 Asp Ser Glu Glu Ser Asp Glu Asp Met Gly Phe Gly Leu Phe Asp
656 305          310          315

```

1130 <210> SEQ ID NO: 22

1131 <211> LENGTH: 2054

1132 <212> TYPE: DNA

1133 <213> ORGANISM: Danio rerio

1135 <220> FEATURE:

1136 <221> NAME/KEY: TATA signal

1137 <222> LOCATION: (1983)..(1989)

1139 <220> FEATURE:

1140 <221> NAME/KEY: enhancer

1141 <222> LOCATION: (142)..(148)

1142 <223> OTHER INFORMATION: E-box, canntg

1144 <220> FEATURE:

1145 <221> NAME/KEY: enhancer

1146 <222> LOCATION: (452)..(457)

1147 <223> OTHER INFORMATION: <223> E-box, canntg

1149 <220> FEATURE:

1150 <221> NAME/KEY: enhancer

1151 <222> LOCATION: (1095)..(1100)

1152 <223> OTHER INFORMATION: E-box, canntg

1154 <220> FEATURE:

1155 <221> NAME/KEY: enhancer

1156 <222> LOCATION: (1278)..(1283)

1157 <223> OTHER INFORMATION: E-box, canntg

1159 <220> FEATURE:

1160 <221> NAME/KEY: enhancer

1161 <222> LOCATION: (1362)..(1367)

1162 <223> OTHER INFORMATION: E-box, canntg

1164 <220> FEATURE:

1165 <221> NAME/KEY: enhancer

1166 <222> LOCATION: (1385)..(1390)

RAW SEQUENCE LISTING

DATE: 03/25/2004

PATENT APPLICATION: US/10/605,708

TIME: 09:07:31

Input Set : D:\GLOF007USC1.txt

Output Set: N:\CRF4\03252004\J605708.raw

```

1167 <223> OTHER INFORMATION: E-box, canntg
1169 <220> FEATURE:
1170 <221> NAME/KEY: enhancer
1171 <222> LOCATION: (523)..(532)
1172 <223> OTHER INFORMATION: Potential MEF2 binding site, yta(w)4tar
1174 <220> FEATURE:
1175 <221> NAME/KEY: enhancer
1176 <222> LOCATION: (606)..(615)
1177 <223> OTHER INFORMATION: Potential MEF2 binding site, yta(w)4tar
1179 <220> FEATURE:
1180 <221> NAME/KEY: enhancer
1181 <222> LOCATION: (697)..(706)
1182 <223> OTHER INFORMATION: Potential MEF2 binding site, yta(w)4tar
1184 <220> FEATURE:
1185 <221> NAME/KEY: enhancer
1186 <222> LOCATION: (1490)..(1499)
1187 <223> OTHER INFORMATION: Potential MEF2 binding site, yta(w)4tar
1189 <220> FEATURE:
1190 <221> NAME/KEY: enhancer
1191 <222> LOCATION: (1640)..(1649)
1192 <223> OTHER INFORMATION: Potential MEF2 binding site, yta(w)4tar
1194 <220> FEATURE:
1195 <221> NAME/KEY: enhancer
1196 <222> LOCATION: (1956)..(1965)
1197 <223> OTHER INFORMATION: Potential MEF2 binding site, yta(w)4tar
1199 <220> FEATURE:
1200 <223> OTHER INFORMATION: Transcription start site at residue 2012
1202 <220> FEATURE:
-> 1203 <221> NAME/KEY: primer bind
1204 <222> LOCATION: (2032)..(2054)
1205 <223> OTHER INFORMATION: M2
1207 <220> FEATURE:
-> 1208 <221> NAME/KEY: misc difference
1209 <222> LOCATION: (2027)..(2054)
1210 <223> OTHER INFORMATION: Identical to the 5' MLC2f cDNA
-> 1211 <409> 22
1212 tgcatgcctg gcaggtccac tctagaggac tactagtcac atgcgattct gaacaatgct 60
1213 ggaatgagcc accaactcat ccagtgtatt accctacact gggaaacacc caaatctgtc 120
1214 tgttatattt gtgcataaac attagattag aagctgtcac tgcggtggta cttttcaaaa 180
-> 1215 ttgataacct aaaagtatat attagtgcct tttagggtact aatatatacc cttgaggttt 240
-> 1216 tcatttggaag aggtaccacc ccagtgcacg aaatctggag cttatttaac aaaataaact 300
-> 1217 tatttatatg ttattgaaaa atattaaata agcaaaacaa tggaaaaaaa ttagttcaaa 360
-> 1218 atttagcttt atttaaattg tttatcttt aatatagctg tttataaat ctgttttgtt 420
-> 1219 actgagagat ggagaaaaat attcattttc ctgtaattat ctgtttttct aggtactgta 480
-> 1220 caagcaggag caaaacaagc cgacagactc gggaaatgcac aacaaactca aggggggcaa 540
-> 1221 gagagcaagg agcgctcaag attgttttagc ctgccttccc aaaaaaaaac tgtcttaagc 600
-> 1222 caaccactca gagggctgta gtgtgctgac cgtgcttgac cacagggcag cttccacaaa 660
-> 1223 gtgaggtcat aggtcgatcg gcagagagag atgggcatgg ccatgtggac ggggtgtgtg 720
-> 1224 actatactag gaaaagcatt aaaacctatt aagacaccag aacgtcctct tatatatcag 780

```

OK
this
is due
to above
error.

Invalid
Numeric
Identifier

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/605,708

DATE: 03/25/2004
TIME: 09:07:31

Input Set : D:\GLOF007USC1.txt
Output Set: N:\CRF4\03252004\J605708.raw

```

W--> 1225      tcattggctc aaaaatctct ggattgaaat atccaacaag taatcctgca agataagcca 840
W--> 1226      ggagggagtt gcgteccctt tagactcagt atgtgattgt atgaagctca aacagtccct 900
W--> 1227      gtggacagct tgaattcaat tcgccacaga ttttatgcag cggatgccca tccagttgca 960
W--> 1228      ttttaaatta atatttttaa taggaagcta tcagtacact ctcaaaaata aatggtcgca 1020
W--> 1229      aggtacatat ttgtacttaa aggggtccata aaaaatttta agagaaacac tttgtacttt 1080
W--> 1230      tattatggac cttaaaggta caaattttta ctcacgcctt ttatttctga gagtgaagct 1140
W--> 1231      atgataacgg tccaaaaact actacacca caaatttata aacaggggaa aatcaagaga 1200
W--> 1232      atttgtaggt tgtaattttt ttgttgcaat caattttgtg actaaaatat tattttaata 1260
W--> 1233      taaatgcacc aaaatacatt gcctatatct aaaaatgggt gtactcaatt actctaagca 1320
W--> 1234      aaataatgct aatcttaaac aattttggaa acaggatatc aaattagtct aaagaaagaa 1380
W--> 1235      aacagtgact gatgaattag acaagaaaaa tattttgggt accacagctg ttccttatgc 1440
W--> 1236      ctcaaatttc tcttcattag ggtccaacat catctaaaaa ctgggaaaaa ggggtaatta 1500
W--> 1237      atggcacctc acagtcactg aagtgaccgg agagagagag agagagagag agtgctgaat 1560
W--> 1238      ggggcacttg aaccgaaatc ttacagcatc ttogattagg gctgatttga aataagggtt 1620
W--> 1239      ccagggcggtg aacaaatatg aacaacataa ccacaggat ctatcaotgc aaccctcccc 1680
W--> 1240      gtattgatct gctgctaata taactttagg ggtacagct cattcatttc aaattgagtt 1740
W--> 1241      tacgtcccca tgtcttatt agacaacgg agacatgcag gccgtgccca tcagtatcag 1800
W--> 1242      attcatecca ttccaagact ccaatagcta tttctgagca ctgtaagatg atagtacatc 1860
W--> 1243      ccagccggtg tccctccatc actttccccc tacctcatag ttttctctct tctctctctg 1920
W--> 1244      gtctgctatt tcccaaacct cacttaaggt tgggtctata attagcaagg ggccttcgct 1980
W--> 1245      agtatataag cccctcaagt acaggacact acgcggtctc agacttctct tcttgatctt 2040
W--> 1246      attgaacttc acac 2054

```

SEQUENCE:

1250 <400>
1250
1250 <210> SEQ ID NO: 23

↑ OK,
this is
due to
above
error
ON previous
page.

VARIABLE LOCATION SUMMARY

PATENT APPLICATION: US/10/605,708

DATE: 03/25/2004

TIME: 09:07:32

Input Set : D:\GLOF007USC1.txt

Output Set: N:\CRF4\03252004\J605708.raw

se of n's or Xaa's (NEW RULES):

se of n's and/or Xaa's have been detected in the Sequence Listing.

se of <220> to <223> is MANDATORY if n's or Xaa's are present.

1 <220> to <223> section, please explain location of n or Xaa, and which
asidue n or Xaa represents.

eq#:17; N Pos. 10

Error Explanation: ✓

<210> 17
<211> 10
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:

Oligonucleotide for linker used in linker-mediated PCR

<220> ~~_____~~ please insert <222> to show "N" location.
<223> n is a dideoxycytidine

<400> 17
gaattcaagn

↑ please see error explanation on
page 9.

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/605,708

DATE: 03/25/2004

TIME: 09:07:32

Input Set : D:\GLOF007USC1.txt

Output Set: N:\CRF4\03252004\J605708.raw

:13 M:270 C: Current Application Number differs, Replaced Application Number
:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date
:31 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
:36 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
:41 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
:52 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:56 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:60 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:65 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:67 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:71 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:73 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1 ✓
:75 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:77 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1 ✓
:79 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:83 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:87 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:91 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:95 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:97 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1 ✓
:99 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:103 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:105 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1 ✓
:107 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:109 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1 ✓
:111 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:115 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:117 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:2 ✓
:119 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:123 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:125 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1 ✓
:127 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:129 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1 ✓
:130 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1
:131 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:133 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:3 ✓
:134 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:2
:135 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:137 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1 ✓
:139 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:141 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1 ✓
:142 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1
:143 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:145 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1 ✓
:147 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:151 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
:153 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1 ✓
:154 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:3

Correct only the
- E's

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/605,708

DATE: 03/25/2004

TIME: 09:07:32

Input Set : D:\GLOF007USC1.txt

Output Set: N:\CRF4\03252004\J605708.raw

155 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
159 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
163 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
167 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
171 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
300 M:252 E: No. of Seq. differs, <211> LENGTH:Input:498 Found:484 SEQ:2 ✓
313 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
318 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
323 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
508 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:5
513 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:5
606 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:5
616 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:6 differs:2 ✓
665 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:7
669 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:7
674 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:7
724 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:8
728 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:8
733 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:8
771 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:9
850 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:11
855 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:11
873 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:12
891 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:13
896 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:13
927 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:15
932 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:15
966 M:258 W: Mandatory Feature missing, <221> Tag not found for SEQ ID#:17
966 M:258 W: Mandatory Feature missing, <222> Tag not found for SEQ ID#:17
966 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17 after pos.:0
1005 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:20
1010 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:20
1015 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:20
1019 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:20
1068 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:20
1136 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:22
1203 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:22
1208 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:22
1211 M:250 E: Invalid Numeric Identifier, INVALID IDENTIFIER ✓
1215 M:259 W: Allowed number of lines exceeded, <223> Other Information:
1216 M:259 W: Allowed number of lines exceeded, <223> Other Information:
1217 M:259 W: Allowed number of lines exceeded, <223> Other Information:
1218 M:259 W: Allowed number of lines exceeded, <223> Other Information:
1219 M:259 W: Allowed number of lines exceeded, <223> Other Information:
1220 M:259 W: Allowed number of lines exceeded, <223> Other Information:
1221 M:259 W: Allowed number of lines exceeded, <223> Other Information:
1222 M:259 W: Allowed number of lines exceeded, <223> Other Information:
1223 M:259 W: Allowed number of lines exceeded, <223> Other Information:
1224 M:259 W: Allowed number of lines exceeded, <223> Other Information:

VERIFICATION SUMMARY

DATE: 03/25/2004

PATENT APPLICATION: US/10/605,708

TIME: 09:07:32

Input Set : D:\GLOF007USC1.txt

Output Set: N:\CRF4\03252004\J605708.raw

225 M:259 W: Allowed number of lines exceeded, <223> Other Information:
226 M:259 W: Allowed number of lines exceeded, <223> Other Information:
227 M:259 W: Allowed number of lines exceeded, <223> Other Information:
228 M:259 W: Allowed number of lines exceeded, <223> Other Information:
229 M:259 W: Allowed number of lines exceeded, <223> Other Information:
230 M:259 W: Allowed number of lines exceeded, <223> Other Information:
231 M:259 W: Allowed number of lines exceeded, <223> Other Information:
232 M:259 W: Allowed number of lines exceeded, <223> Other Information:
233 M:259 W: Allowed number of lines exceeded, <223> Other Information:
234 M:259 W: Allowed number of lines exceeded, <223> Other Information:
235 M:259 W: Allowed number of lines exceeded, <223> Other Information:
236 M:259 W: Allowed number of lines exceeded, <223> Other Information:
237 M:259 W: Allowed number of lines exceeded, <223> Other Information:
238 M:259 W: Allowed number of lines exceeded, <223> Other Information:
239 M:259 W: Allowed number of lines exceeded, <223> Other Information:
240 M:259 W: Allowed number of lines exceeded, <223> Other Information:
241 M:259 W: Allowed number of lines exceeded, <223> Other Information:
242 M:259 W: Allowed number of lines exceeded, <223> Other Information:
243 M:259 W: Allowed number of lines exceeded, <223> Other Information:
244 M:259 W: Allowed number of lines exceeded, <223> Other Information:
245 M:259 W: Allowed number of lines exceeded, <223> Other Information:
246 M:259 W: Allowed number of lines exceeded, <223> Other Information:
250 M:200 E: Mandatory Header Field missing, <400> is required.
250 M:301 E: (44) No Sequence Data was Shown, SEQ ID:22
250 M:252 E: No. of Seq. differs, <211> LENGTH: Input:2054 Found:0 SEQ:22
250 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQ ID NO:23
7 M:203 E: No. of Seq. differs, <160> Number Of Sequences: Input (24) Counted (23)